Introduction

- Belief Propagation Techniques uses the degree of a person's belief that an event will occur,rather than the actual probability of the event will occur.
- Belief probabilities are properties of person's belief not the event.
- Belief Propagation Technique is used to perform inference on graphical model such as factor graphs which calculates marginal distribution for each unobserved node conditional on observed node.
- Belief Propagation Techniques are used for performing inference on graphical models such as Bayesian Network and Markov Random Field

Outline of the seminar

1. **Part 1:Concepts Related to Belief Propagation Techniques**
2. **Part 2:Applications of Belief Propagation Techniques**

Outline-Part 1

- **Probabilistic Formulation**
- **Probability Model**
- **Probability Language**
- **Graphical Representation**
- **Bayesian Belief Network**
- **Markov Random Field**

Probabilistic Formulation

- Bayesian Methods provides reasoning about partial beliefs under condition of uncertainty
- Beliefs measures obey the three basic axioms of probability theory.

$$0 \leq P(A) \leq 1 \tag{1}$$

$$P(SurePropositions) = 1 \tag{2}$$

$$P(AorB) = P(A) + P(B) \tag{3}$$

if A and B are mutually exclusive

Probabilistic Formulation

- The basic expressions in the Bayesian formalism are statement about
  1. Conditional probabilities
  2. Product rule is so called **Chain Rule Formula**.
- $$P(A) = \sum \{P(A \setminus B_i) \times P(B_i)\} \tag{4}$$
- The probability of any event A computed by conditioning it on any set of exhaustive and mutually exclusive events $B_i, i = 1, 2, ...n$
- This decomposition provides the basics for the hypothetical or assumption based reasoning in the bayesian formalism.

Probabilistic Formulation:product rule or Chain Rule Formula

- It states that if a set of n events $(E_1, E_2, ...E_n)$
- probability of joint event $(E_1, E_2, ...E_n)$ can be written as a product of n conditional probabilities
- $$P(E_1, E_2, ...E_n) = P(E_n|E_{n-1}, ...E_2, E_1)......P(E_2|E_1)P(E_1) \tag{5}$$
- This product can be derived by repeated application of equation 6 in any convenient order.
- $$P(A, B) = P(A|B)P(B) \tag{6}$$

Probability Model

- It is an encoding of the probabilistic information in some domain of interest
- It is formed sentence,statement or proposition which can be calculated in accordance with axioms of probability theory.
- These sentences are represented by interrelating random or stochastic variables with in model.
- The probability model (M) is defined by the joint probability of all its variables
- Each variable may take any one of a set of mutually exclusive or collectively exhaustive states.

Probability Language

- It is suited to reasoning under uncertainty and provides suitable frame work for processing the uncertainty relationship between variables of a probabilistic model.
- Through axiomatic basis and provide a convenient mechanism for presenting uncertain results.
- The basic four primitives of probability language are
  1. Likelihood
  2. Conditioning
  3. Relevance
  4. Causation

Probability Language

- **A likelihood** of event is measure of how likely or probable that event will occur ie.chance of occurrence.
- **A conditioning :**An event is conditional when it changes by the knowledge of second event states.
- **Relevance:** Two events are relevance,when common sequences is observed.ie event A and event B are relevant if both are causes of event C.
- **Causation** conveys a pattern of dependence between events.

Graphical Representation

- A probabilistic model is dependency model in which relationship between each of the variable is captured.
- Graphs may be used to represent dependency model.Dependency Model M comprising the variables U represented by graph.
- A graph is denoted by G(V,E) is set of nodes or vertices V connected by the set of arcs or edge E.
- The set of arc E represents conditional dependencies between these variables.
- ie.Joint probability function of M is encoded in E,Nodes of the graph corresponds to variables in dependency model

Graphical Representation

- Graphical representation for probabilistic model are two types
- **Undirected graph and Directed graph.**
- Undirected graph had no explicit direction,an arc of influence between connected nodes.
- Examples of undirected graph is Markov Random Fields.
- In directed graph arcs are either unidirectional or bidirectional directed arc provide the mechanism for representing causation
- Example for Directed graph is Bayesian network

Bayesian Belief Network

- The Bayesian Belief Network(BBN) determines the state probabilities of each node or variable from predetermined conditional and prior probabilities.
- The Direct Acyclic Graph G(U,E)of the probability model M represents probability distribution P(U) ,where U represents set of all variables in M
- Let X_1,X_2,........X_n are Variables in probability distribution P(U)
- Direct Acyclic Graph(DAG) in which minimal set of variables are designed as parent of each variable is X_i such that

$$P(X_i/W); W \in X_1, X_2.....X_{i-1} \tag{7}$$

- The above equation is BBN of that probability distribution

Bayesian Belief Network

$$P(a/b)P(b) = P(a,b) \tag{8}$$

In real world all events are conditioned by some context C=c

$$P(a/b,c)P(b/c) = P(a,b/c) \tag{9}$$

The Baye's rule also conditioned on

$$P(b/a,c) = P(a/b,c)P(b/c)/P(a/c) \tag{10}$$

- P(b/a,c) is posterior probability of b
- P(a/b,c) liklihood probability
- P(a/c) Normalized factor
- P(b/c) Priori probability of b

Bayesian Belief Network

Normalised factor for continuous and discrete distributions are given by

$$\int_b P(a/b, c)p(b/c)db \tag{11}$$

$$\sum_b P(a/b, c)p(b/c) \tag{12}$$

Markov Random Field

The following points to consider while applying Markov Random Field to the Belief propagation Technique

- Markov Random Field expressed in terms of conditional independence of its non neighbors,once values of neighbors are known.
- Assigning of weights to the links of the graph must be handled with caution.
- The weights are to be used translating evidential data in to meaning full probabilistic inferences such probabilistic model is both consistent and complete.
- Consistency guarantees that,it don't over load the graph with too many parameters.

Markov Random Field

- For unidirected graph G ,nodes are always adjust to each other.
- For each group of nodes known as associate ,assign a nonnegative compatibility function $g_i(c_i)$,which measures the relative degree of compatibility associated with each value assignment c_i to the variable.
- Form the product $\prod_i g_i(c_i)$ of the compatibility functions over all associates.
- Normalize the product of all possible value combinations of the variables of the system

Markov Random Field

$$P(x_i, ..., x_n) = K \prod_i g_i(c_i) \tag{13}$$

where

$$K = 1 \div [\sum_{x_1,...x_n} \prod_i g_i(c_i)] \tag{14}$$

The normalized product P in equation 13 is a joint distribution that consists of all conditional independencies of graph G.

Outline:Part-2

1. Efficient Belief Propagation for Early Vision.
2. Efficient Loopy Belief Propagation using the Four Color Theorem
3. Markov Network-based Unified Classifier for Face Identification
4. Image Completion Using Efficient Belief Propagation Via Priority Scheduling and Dynamic Pruning
5. Low Memory Cost Block-Based Belief Propagation For Stereo Correspondence
6. Task Parallel Implementation of Belief Propagation in Factor Graphs
7. Hardware-Efficient Belief Propagation
8. PatchMatch Belief Propagation for Correspondence Field Estimation
9. Learning continuous time Bayesian network classifiers

1.Efficient Belief Propagation for Early Vision

- Early vision problems such as stereo and image restoration
- Inference algorithms based on graph cuts and belief propagation give accurate results, but too slow for practical use.
- Markov random field models are used in this application.
- In this application some of the techniques are used that substantially improve the running time of the loopy belief propagation .

1.Efficient Belief Propagation for Early Vision

1. One of the techniques reduces the complexity of the inference algorithm to be linear rather than quadratic by fixing the number of possible labels for each pixel, which is important for problems such as image restoration that have a large label set.
2. second technique speeds up and reduces the memory requirements of belief propagation on grid graphs.
3. A third technique is a multi-grid method that makes it possible to obtain good results with a small fixed number of message passing iterations, independent of the size of the input images.

1.Efficient Belief Propagation for Early Vision

Results

- Taken together these techniques speed up the standard algorithm by several orders of magnitude and results obtained are.
- The time necessary to compute a single message update from O (k^2) to O (k), where k is the number of possible labels for each pixel for the max-product formulation ie.linear rather than quadratic is possible.

2.Efficient Loopy Belief Propagation using the Four Color Theorem

- Recent work on early vision such as image segmentation, image denoising, stereo matching, and optical flow uses Markov Random Fields.
- Although this formulation yields an NP-hard energy minimization problem, good heuristics have been developed based on graph cuts and belief propagation.
- Both approaches still require tens of seconds to solve stereo problems on recent PCs. Such running times are impractical for optical flow and many image segmentation and denoising problems

2.Efficient Loopy Belief Propagation using the Four Color Theorem

- To reduce the computational complexity of belief propagation can be achieved by applying Four Color Theorem (FCT) which limits the maximum number of labels in image segmentation to four.
- Four Color Theorem (FCT) states that when an image seen as a planar graph is segmented into contiguous regions, there are only four colors to be assigned to each pixel/node for all other segments to be surrounded by different colors .

2.Efficient Loopy Belief Propagation using the Four Color Theorem

- In the case of Belief Propagation (BP), a key reason for its slow performance is that the algorithm complexity is proportional to both the number of pixels in the image, and the number of labels in the underlying image segmentation which is typically high. If limit the number of labels, its speed performance should improve greatly.
- By modifying the propagation algorithms like can using a low number of placeholder labels, that can reuse for non-adjacent segments. These placeholder labels can then be replaced by the full set of actual labels.

2.Efficient Loopy Belief Propagation using the Four Color Theorem

- Since image segments form a planar graph Four Color Theorem (FCT)can be applied to fix the four placeholder labels
- A fast segmentation through the placeholder labels and a fine grained labeling through the actual labels provides a joint optimization process
- The computational time is basically dependent on the number of placeholder rather than actual labels.

2.Efficient Loopy Belief Propagation using the Four Color Theorem

Results

- Four-Color Theorem (FCT)based on the max-product belief propagation technique can be used in early computer vision for solving MRF problems where an energy is to be minimized.
- The Methods used can improve either the speed for large images and/or large label sets
- FCT principle is difficult to apply where label set is discrete in the case for stereo matching and optical flow where the disparity cost function takes discrete and unrelated values.This causes slower convergence but FCT methods can be used to solve the above mentioned problems.

3.Markov Network-based Unified Classifier for Face Identification

- It is a one-to-many identification problem and has many applications such as searching for similar face images in a database and face tagging in images and videos.
- Recent successful face recognition methods,classifiers are used where similar scores are merged with the predefined parameters.
- The parameter comes from the training database and it is not the best choice when the input image has different conditions.
- These methods lead to good accuracy in face verification, but there is no specific framework for the one-to-many identification problem.

3.Markov Network-based Unified Classifier for Face Identification

- In this paper, a novel recognition framework for the one-to-many identification is designed, and the simple concept is illustrated in Figure 1.
- The multiple classifiers have complementary characteristics, unify the multiple classifiers based on a Markov network as shown in Figure 2

3.Markov Network-based Unified Classifier for Face Identification

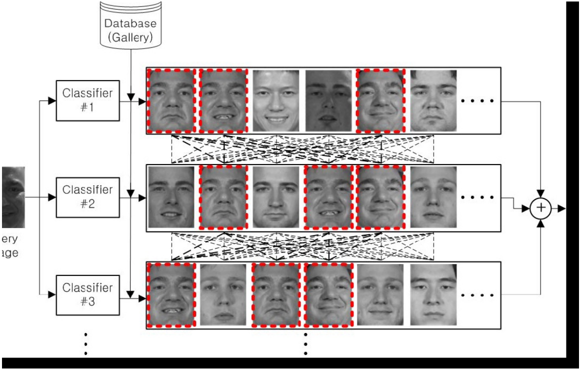

Figure: 1.One-to-Many identification

3.Markov Network-based Unified Classifier for Face Identification

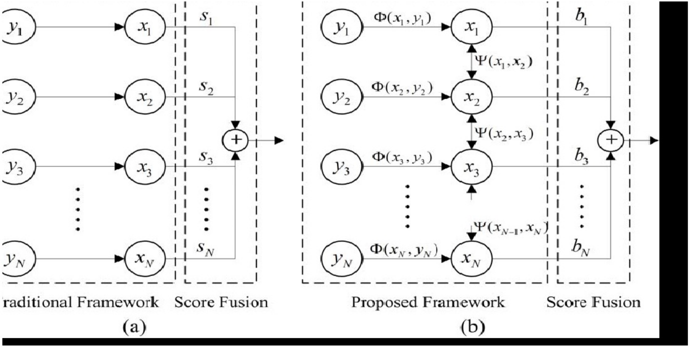

Figure: 2.Traditional and proposed Frame Work

3.Markov Network-based Unified Classifier for Face Identification

1. steps to find marginal probability by using Markov Fields for classifiers are as follows
2. one node of a Markov network to each classifier,Nodes are connected by lines, which represents the statistical dependencies
3. for an observation node, extract a feature from a query image using the corresponding classifier.
4. At its paired hidden node, the first retrieve n similar gallery samples from the database.

3.Markov Network-based Unified Classifier for Face Identification

1. The multiple classifiers have their own lists of retrieved gallery images, which are not identical in general, thereby complementing the neighbor classifiers.
2. the hidden nodes are connected by the network lines,
3. the relationship between nodes is by selecting similarity scores between the neighbor classifiers,then scores are calculated by concatenating two gallery features of the neighbor classifiers.
4. The posterior probability at each hidden node is calculated by the belief-propagation algorithm. Finally, marginal probability for a score value at each classifier is calculated

3.Markov Network-based Unified Classifier for Face Identification

- The results obtained are for face recognition particularly for the one-to-many identification task, based on multiple classifiers gallery connected by a Markov network.
- The Markov network probabilistically models the relationships between a query images and between neighboring gallery images.
- The observation-hidden node pair retrieve the similar gallery images from the database image.
- The similarities between the retrieved gallery images gives the statistical dependency between the hidden nodes.
- Hence results obtained can be viewed as clustering-based face recognition.

4.Image Completion Using Efficient Belief Propagation Via Priority Scheduling and Dynamic Pruning

Any algorithm that is designed to solve the image completion problem should have the following characteristics:

1. It should be able to successfully complete complex natural images
2. It should also be able to handle incomplete images with (possibly) large missing parts
3. All these should take place in a fully automatic manner, i.e., without intervention from the user.

4.Image Completion

Figure: Object removal

Object removal is just one of the many cases where image completion needs to be applied.

4.Image Completion Using Efficient Belief Propagation Via Priority Scheduling and Dynamic Pruning

Three main approaches for dealing with the image completion problem

1. Statistical-based methods
2. Partial differential equations (PDE)based methods
3. Exemplar-based methods

4.Image Completion Using Efficient Belief Propagation Via Priority Scheduling and Dynamic Pruning

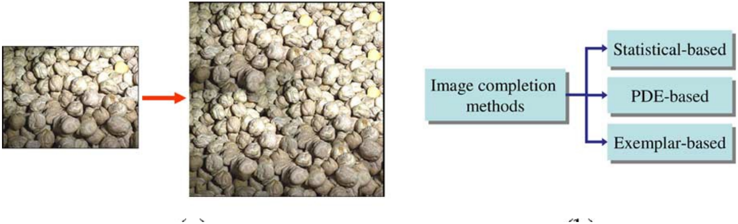

Figure: Image Completion

4.Image Completion Using Efficient Belief Propagation Via Priority Scheduling and Dynamic Pruning

- Exemplar-based techniques fill the unknown region simply by copying content from the observed part of the image.
- All exemplar-based techniques for texture synthesis were either pixel-based or patch-based, meaning that the final texture was synthesized one pixel, or one patch at a time
- A new exemplar-based framework which treats image completion, texture synthesis, and image inpainting in a unified manner.
- All image-editing tasks in the form of a discrete global optimization which is used to avoid visually inconsistent results.

4.Image Completion Using Efficient Belief Propagation Via Priority Scheduling and Dynamic Pruning

- Novel optimization scheme, called priority belief propagation (BP) is used which carries two very important extensions over the standard BP algorithm
  1. **Priority-based message scheduling**
  2. **Dynamic label pruning**
- As one of major limitation of the BP algorithm is its inefficiency in handling MRFs with very large discrete state spaces is considered to resolve by these extension techniques.

4.Image Completion Using Efficient Belief Propagation Via Priority Scheduling and Dynamic Pruning

- Priority Belief Propagation can be used for other types of completion problems such as video completion or geometric completion,constrained texture synthesis
- Priority-BP algorithm, which is a generic MRF optimization scheme can be used to other labeling problems also.

5.Low Memory Cost Block-Based Belief Propagation For Stereo Correspondence

- Stereo correspondence is used in computer vision to find the depth among the cameras and objects.
- The depth inference problem could be further transformed to a disparity inference problem by assuming that the cameras and objects are under epipolar geometry.
- The inferred disparity information could be widely applied to tracking, surveillance system, and multiview video coding.

5.Low Memory Cost Block-Based Belief Propagation For Stereo Correspondence

The stereo matching algorithms can be roughly divided into two categories.

1. **Local approaches**
2. **Global approaches**

5.Low Memory Cost Block-Based Belief Propagation For Stereo Correspondence

- Local approaches select disparities of image pixels using the information in a window. Therefore local approaches are faster than global approaches.
- It results in poor accuracy since the local approaches could not deal with textureless regions and occluded regions well due to the insufficient information in window.
- On the other hand, global approaches can handle the textureless and occluded regions well by formulating disparity inference as an energy minimization problem

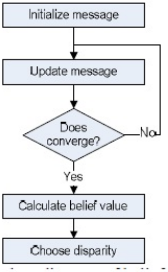

Figure: Flow diagram of Belief Propagation

5.Low Memory Cost Block-Based Belief Propagation For Stereo Correspondence

- The BP algorithms construct 2-D graph structures with nodes representing all the pixels in the disparity images to find the disparity map with energy closer to the global minima
- The vast number of nodes in the 2-D graph result in extremely high computation complexity, thereby rendering 2-D optimization is too difficult to be directly implemented for real-time application.
- A block-based BP algorithm that directly partitions an image into separated independent blocks. Thus, can reduce the memory size significantly due to block based computation.
- The independent blocks also enable parallel computation by multiple computation units. Moreover earlier convergence for each block can also improve the long running time

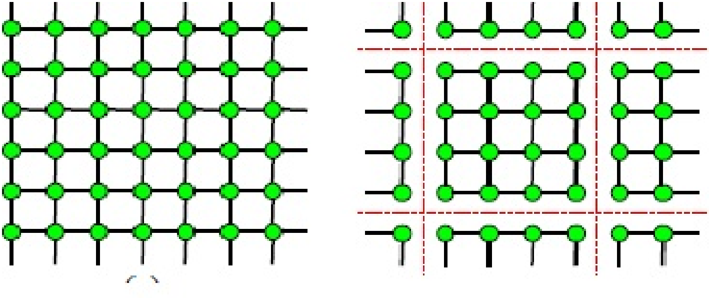

Figure: Graph of typical and block based BP

2D Graph Model

5.Low Memory Cost Block-Based Belief Propagation For Stereo Correspondence

Results

- A new stereo matching algorithm partitions an image to block and optimizes with belief propagation technique.This method reduces memory storage size by 99 percentage with good performance.
- To enhance the interaction between neighboring blocks such that the independent block could extract useful information from neighboring finished processing blocks are possible with block belief propagation.

6.Task Parallel Implementation Of Belief Propagation in Factor Graphs

- Graphical models are essential tools for probabilistic reasoning. Factor graphs are emerged as a unified model of directed graphs (e.g. Bayesian networks) and undirected graphs (e.g. Markov networks).
- A factor graph represents a joint probability distribution ie.written as a product of factors, each involving a subset of random variables.
- Factor graphs have found applications in Image processing,Bioinformatics and Error-control decoding used in digital communications.

6.Task Parallel Implementation Of Belief Propagation In Factor Graphs

Relation between Factor graphs and Belief Propagation.

- Inference is a problem of computing posterior probability distribution of certain variables given some value as observed or evidence variables.
- In factor graphs, inference proceeds with the belief propagation algorithm .
- Belief propagation is a process of passing messages along the edges of a graph.
- Processing each message requires a set of operations with respect to the probability distribution of the random variables in a graph.
- Such distribution is represented by potential tables.
- The complexity of belief propagation increases dramatically as the number of states of variables and node degrees of a graph increase.

6.Task Parallel Implementation Of Belief Propagation in Factor Graphs

Results

- Belief propagation must be performed in real time.Many parallel techniques have been proposed for belief propagation in factor graphs.
- Parallelizing belief propagation in acyclic factor graphs still remains a challenging problem due to the precedence constraints among the nodes in the graphs.
- Task scheduling is used in parallel computing is an efficient tool for linear algebra problem on general-purpose multi-core processors.
- The methods used for implementation of Belief Propagation using factor graphs are task dependency graph by using a dynamic task scheduler.

7.Hardware-Efficient Belief Propagation

- The success of BP is due to its regularity and simplicity.
- It uses a simple message update process to iteratively refine the beliefs of labels for each node.
- A message sent from one node to another is updated according to neighboring messages and local energy functions.
- Loopy belief propagation (BP)is an effective solution for assigning labels to the nodes of a graphical model such as the Markov random field (MRF).

7.Hardware-Efficient Belief Propagation

- BP algorithms generally require a great amount of memory for storing the messages, typically on the order of tens to hundreds times larger than the input data.
- Since each message is processed hundreds of times, the saving/loading of messages consumes considerable bandwidth.
- Although BP may work on high-end platforms such as desktops, it cannot be applied to most consumer electronic devices that have limited memory, computational power, and energy.
- It is difficult to utilize hardware parallelism to accelerate BP

7.Hardware-Efficient Belief Propagation

The two techniques are used for sequential procedure to accelerate BP

- The first one is tile-based BP splits the Markov random field (MRF) into many tiles and only stores the messages across the neighboring tiles. The memory and bandwidth required by this technique is only a fraction of the ordinary BP algorithms.
- Second technique is that fast message construction technique is based on the observation that many hypotheses used to construct the mesasages are repetitive.therefore, they only need to be computed once which reduces running time of the alogorithm

7.Hardware-Efficient Belief Propagation

Results

- These techniques can be realized in both hardware and software.
- A software reference implementation compatible to the Middlebury MRF library which is available online.
- First hardware is a very large scale integration (VLSI) circuit and the second one a graphic processing unit (GPU) program are analyzed
- **A tile-based message passing and fast message construction algorithm** greatly reduced the memory, bandwidth, and computational costs of BP and enabled the parallel processing. With these two techniques BP can be more suitable for low-cost and power limited consumer electronics.

8.PMBP: PatchMatch Belief Propagation for Correspondence Field Estimation

Patch Match is a simple, yet very powerful and successful method for optimizing Continuous labelling problems. The two main approaches used are

- The update of the solution space by sampling
- The use of the spatial neighborhood to propagate samples.

These approaches are related to steps in a specific form of belief Propagation in the continuous space, called Particle Belief Propagation (PBP)

8.PMBP: PatchMatch Belief Propagation for Correspondence Field Estimation

- The two approaches used in this research yields a new algorithm called Patch Match Belief Propagation for Correspondence Field estimation (PMBP),which is more accurate than Patch Match and orders of magnitude faster than PBP.
- The link between the popular PatchMatch method and the very well-known Belief propagation algorithm.
- These approaches can be used as both in terms of applications, such as optical flow, as well as algorithms such as different forms of message passing e.g. Treereweighted

9.Learning continuous time Bayesian network classifiers

- Continuous time Bayesian network classifiers are designed for analyzing multivariate streaming data when time duration of event matters.
- The continuous time Bayesian network classifiers are considered in the case where complete data is available.
- Conditional log-likelihood scoring is developed for structural learning on continuous time Bayesian network classifiers.
- Results show that conditional log-likelihood scoring combined with Bayesian parameter estimation outperforms marginal log-likelihood scoring.
- Conditional log-likelihood scoring becomes even more effective when the amount of available data is limited.

9.Learning continuous time Bayesian network classifiers

- Conditional log-likelihood scoring function is used to learn continuous time Bayesian network classifiers from multivariate streaming data.
- Same function can be used for classifying multivariate trajectories in the case where the class is static .
- The quality of the classification performances also suggests to extend the Continuous Time Bayesian Network Classifiers to the clustering problem.

Conclusion

- The graphical representation or models for multidimentional probability distributions such as Markov Random Fields and Bayesian Networks are used for Belief Propagation Techniques.
- Some of the applications to optimize Markov Random Fields to speed up the Belief Propagation are used for face recognition, for image completion and low level vision problem,
- To reduce memory size and computation time block based Belief Propagation algorithm ,tile based Belief Propagation and fast construction techniques are used.

References1

1. Judea Pearl, **Probabilistic Reasoning in Intelligent Systems**, Kaufmann Morgann Publishers, Second Revised And updated Edition,San Francisco, California, 1998.
2. Yu-Chang And Nelson Chang,**Low Memory Cost Block-Based Belief Propagation For Stereo Correspondence**, 1-4244-1017-7187,IEEE Transisition
3. F.Besse,C.Rother,A.Fitzgibbon,J.Kautz **Pmbp:Patchmatch Belief Propagation For Correspondence Field Estimation**, University College Of London U.K.,Microsoft Research Cambridge ,2012
4. Nam Ma,Yinglong Xia,Viktorprasanna **Task Parallel Implementation Of Belief Propagation In Factor Graphs** This Research Was Suppored By U.S.National Foundation Under Grant No.Cns-101880

References2

1. Wonjun Hwang Kyungshik,Roh Junmo Kim **Markov Network-Based Unified Classifier For Face Identification** This ICCV 2013 Paper is Open Access Available In IEEE Xplore
2. Radu Timofte,Lucvan Gool **Efficient Loopy Belief Propagation Using The Four Color Theorem** Isics,Esat-Psi/Iminds V,Belgium,Computer Vision Lab,Zurich,Switzerland
3. Nikos Komadakis,Georgios Tziritas **Image Completion Using Efficient Belief Propagation Via Priority Scheduling And Dynamic Pruning** IEEE Transactions In Image Processing ,Volume;16,No.11,November 2007

References3

1. Chia-Kai Liang,Chao-Chung Cheng,Yen-Cheieh Lai **Hardware-Efficient Belief Propagation** IEEE Transactions On Circuits And Systems For Video Technology Vol;21,No.5,May 2011
2. Daniele Codecasa,Fabiostella **Learning Continuous Time Bayesian Network Classifiers** International Journal Of Approximate Reasoning,55(2014) 1728-1746
3. Pedro F.Felzenszwalb,Daniel P.Huttenlocher **Efficient Belief Propagation For Early Vision** International Journal Of Computer Vision (Springer Science)70(1),41-54,2006
4. Jianguo Ding **Probabilistic Inferences In Bayesian Networks** International Journal Of Computer Science,Arxiv:1011.0935V2,5 November 2010

Thank You...